

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-292971

(43)公開日 平成9年(1997)11月11日

(51)Int.Cl. ^a	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 3/16	3 2 0		G 0 6 F 3/16	3 2 0 H
17/21			G 1 0 L 3/00	R
17/28				5 1 1
G 1 0 L 3/00				5 5 1 C
	5 1 1		G 0 6 F 15/20	5 6 8 Z

審査請求 未請求 請求項の数14 O L (全 14 頁) 最終頁に続く

(21)出願番号 特願平8-106746

(22)出願日 平成8年(1996)4月26日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 石井 和夫

東京都品川区北品川6丁目7番35号 ソニ

一株式会社内

(72)発明者 藤村 聡

東京都品川区北品川6丁目7番35号 ソニ

一株式会社内

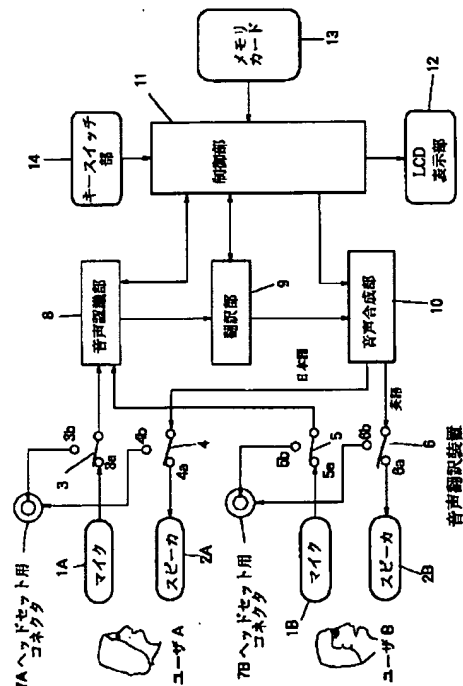
(74)代理人 弁理士 稲本 義雄

(54)【発明の名称】 翻訳装置

(57)【要約】

【課題】 翻訳装置の利便性を向上させる。

【解決手段】 ユーザAにより日本語で発話が行なされると、その音声は、マイク1Aを介して音声認識部8に供給される。マイク1Aから入力された音声は日本語とみなされ、音声認識部8で音声認識され、さらに翻訳部9で英語に翻訳される。そして、その結果得られる英語の翻訳文に対応する合成音が、音声合成部10で生成され、ユーザBが使用するスピーカ2Bに供給されて出力される。一方、ユーザBにより英語で発話が行なされると、その音声は、マイク1Bを介して音声認識部8に供給される。マイク1Bから入力された音声は英語とみなされ、音声認識部8で音声認識され、さらに翻訳部9で日本語に翻訳される。そして、その結果得られる日本語の翻訳文に対応する合成音が、音声合成部10で生成され、ユーザAが使用するスピーカ2Aに供給されて出力される。



【特許請求の範囲】

【請求項 1】 音声を入力する入力手段と、前記入力手段に入力された前記音声を音声認識する音声認識手段と、前記音声認識手段の音声認識結果を、前記入力手段に入力された音声の言語と異なる言語に翻訳する翻訳手段と、前記翻訳手段の翻訳結果を出力する出力手段とを備える翻訳装置であって、前記入力手段は、複数設けられていることを特徴とする翻訳装置。

【請求項 2】 前記入力手段は、所定の方向に指向性を有するマイクロフォンであることを特徴とする請求項 1 に記載の翻訳装置。

【請求項 3】 前記翻訳手段の翻訳結果に対応する合成音を生成する生成手段をさらに備え、前記出力手段は、前記生成手段により生成される前記合成音を出力することを特徴とする請求項 1 に記載の翻訳装置。

【請求項 4】 前記出力手段は、複数設けられていることを特徴とする請求項 3 に記載の翻訳装置。

【請求項 5】 前記入力手段および出力手段は、ユーザの頭部に装着可能になされていることを特徴とする請求項 4 に記載の翻訳装置。

【請求項 6】 前記複数の入力手段および出力手段のうちの少なくとも 1 組は、外部から接続可能になされていることを特徴とする請求項 4 に記載の翻訳装置。

【請求項 7】 所定の端末との通信を制御する通信制御手段をさらに備えることを特徴とする請求項 1 に記載の翻訳装置。

【請求項 8】 前記音声認識手段は、前記通信制御手段により受信される音声も音声認識し、前記翻訳手段は、前記通信制御手段により受信された音声の、前記音声認識手段による音声認識結果も、その音声の言語と異なる言語に翻訳することを特徴とする請求項 7 に記載の翻訳装置。

【請求項 9】 情報を再生する再生手段をさらに備え、前記出力手段は、前記再生手段により再生される前記情報も出力することを特徴とする請求項 1 に記載の翻訳装置。

【請求項 10】 音声を入力する入力手段と、所定の情報に対応する合成音を出力する出力手段とを有する入出力部と、前記入力手段に入力された前記音声を音声認識する音声認識手段と、前記音声認識手段の音声認識結果を、前記入力手段に入力された音声の言語と異なる言語に翻訳する翻訳手段と、前記翻訳手段により出力される翻訳結果に対応する合成音を生成し、前記出力手段に供給する生成手段とを有す

る本体とを備え、

前記入出力部と本体とは、分離して設けられていることを特徴とする翻訳装置。

【請求項 11】 前記入出力部と本体との間では、データの送受が無線で行われることを特徴とする請求項 10 に記載の翻訳装置。

【請求項 12】 前記入力手段は、筆記具に取り付け可能になされていることを特徴とする請求項 10 に記載の翻訳装置。

【請求項 13】 前記入出力部は、ユーザの頭部に装着可能になされていることを特徴とする請求項 10 に記載の翻訳装置。

【請求項 14】 前記出力手段は、所定の接続線を介して前記本体と接続されており、

前記入力手段は、前記所定の接続線の途中に取り付けられていることを特徴とする請求項 10 に記載の翻訳装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、翻訳装置に関する。特に、例えば携帯に便利な翻訳装置に関する。

【0002】

【従来の技術】図 11 は、従来の携帯型の翻訳装置（音声翻訳機）の一例の構成を示している。この翻訳装置においては、例えば、日本語の入力文を、音声により入力すると、その入力文を英語などに翻訳した翻訳文に対応する合成音出力されるとともに、その逆に、英語の入力文を、音声により入力すると、それを日本語に翻訳した翻訳文に対応する合成音出力されるようになっている。

【0003】即ち、マイク（マイクロフォン）1A は、そこに音声が入力されると、その音声を、電気信号としての音声信号に変換し、音声認識部 8 に出力するようになされている。スピーカ 2A は、音声合成部 10 より供給される合成音を出力するようになされている。

【0004】音声認識部 8 は、制御部 11 の制御の下、マイク 1A からの音声信号に基づき、例えば HMM (Hidden Markov Model) 法その他の所定の音声認識アルゴリズムにしたがって音声認識を行い、その音声認識結果として、マイク 1A に入力された音声に対応する、例えばテキストデータなどを出力するようになされている。このテキストデータは、音声認識部 8 から翻訳部 9 に供給されるようになっている。

【0005】翻訳部 9 は、制御部 11 の制御の下、音声認識部 8 からの所定の言語のテキストデータを、それと異なる言語に翻訳し、その結果得られる翻訳文（テキストデータなど）を、音声合成部 10 に出力するようになされている。音声合成部 10 は、制御部 11 の制御にしたがい、翻訳部 9 からの翻訳文に対応する合成音を生成し、スピーカ 2A に供給するようになされている。

【0006】制御部11は、キースイッチ部14の操作に対応して、音声認識部8、翻訳部9、および音声合成部10を制御するようになされている。さらに、制御部11は、LCD（液晶ディスプレイ）表示部12を制御することにより、所定の情報を表示させるようになされている。また、制御部11には、メモリカード13が装着可能になされており、制御部11は、このメモリカード13に記憶されたデータを、音声認識部8、翻訳部9、または音声合成部10に供給することで、それぞれに音声認識処理、翻訳処理、または音声合成処理を行わせるようになされている。

【0007】メモリカード13には、所定の言語L1およびそれと異なる言語L2による音声を音声認識するために必要なデータ、言語L1とL2との間の翻訳を行うために必要なデータ、並びに言語L1およびL2による合成音を生成するために必要なデータ（以下、適宜、これらのデータすべてを含めて言語データという）が記憶されている。従って、メモリカード13を他のものと交換することによって、種々の言語の音声認識および音声合成、並びに種々の言語間の翻訳を行うことができるようになされている。

【0008】キースイッチ部14は、制御部11に対し、所定の指令を与えるときに操作されるようになされている。

【0009】次に、例えば、日本人のユーザ（日本語の話者）Aと、アメリカ人のユーザ（英語の話者）Bとが、この翻訳装置を用いて会話を行う場合を例に、その動作について説明する。従って、メモリカード13には、日本語および英語の言語データが記憶されているものとする。

【0010】まず、ユーザAが、ユーザBに対して発話を行う場合においては、例えば、ユーザAは、キースイッチ部14（言語選択ボタン）を、日本語の入力文を英語に翻訳するように操作する。この操作に対応する操作信号は、制御部11において受信され、制御部11は、この操作信号を受信すると、音声認識部8を日本語の音声を音声認識するように、翻訳部9を日本語を英語に翻訳するように、音声合成部10を英語の合成音を生成するように、それぞれ制御する。

【0011】その後、ユーザAが、マイク1Aに対して、日本語による発話を行うと、その音声は、マイク1Aを介することにより音声信号とされ、音声認識部8に供給される。音声認識部8では、マイク1Aからの音声信号に基づいて音声認識が行われ、これにより、マイク1Aに入力された日本語の音声認識される。そして、その音声認識の結果得られるテキストデータは、翻訳部9に供給される。

【0012】翻訳部9では、音声認識部8からの日本語のテキストデータが英語に翻訳され、その結果得られる英語の翻訳文が音声合成部10に供給される。音声合成

部10では、翻訳部9からの翻訳文に対応した合成音（英語の合成音）が生成され、スピーカ2Aに供給される。これにより、スピーカ2Aからは、ユーザAが発した日本語の音声を英語に翻訳した合成音が出力される。

【0013】次に、ユーザBが、ユーザAに対して発話を行う場合においては、例えば、ユーザBは、キースイッチ部14を、英語の入力文を日本語に翻訳するように操作する。この操作に対応する操作信号は、制御部11において受信され、制御部11は、この操作信号を受信すると、音声認識部8を英語の音声を音声認識するように、翻訳部9を英語を日本語に翻訳するように、音声合成部10を日本語の合成音を生成するように、それぞれ制御する。

【0014】その後、ユーザBが、マイク1Aに対して、英語による発話を行うと、その音声は、マイク1Aを介することにより音声信号とされ、音声認識部8に供給される。音声認識部8では、マイク1Aからの音声信号に基づいて音声認識が行われ、これにより、マイク1Aに入力された英語の音声認識される。そして、その音声認識の結果得られるテキストデータは、翻訳部9に供給される。

【0015】翻訳部9では、音声認識部8からの英語のテキストデータが日本語に翻訳され、その結果得られる日本語の翻訳文が音声合成部10に供給される。音声合成部10では、翻訳部9からの翻訳文に対応した合成音（日本語の合成音）が生成され、スピーカ2Aに供給される。これにより、スピーカ2Aからは、ユーザBが発した英語の音声を日本語に翻訳した合成音が出力される。

【0016】なお、音声認識部8において得られる音声の認識結果および翻訳部9において得られる翻訳文は、制御部11にも供給されるようになされており、制御部11は、この認識結果および翻訳文を受信すると、LCD表示部12に供給して表示させるようになされている。従って、この場合、LCD表示部12では、マイク1Aに入力された音声認識結果と、それを翻訳した翻訳結果とが同時に表示される。

【0017】次に、図12は、従来の携帯型の翻訳装置の他の一例の構成を示している。液晶表示部（LCD（液晶ディスプレイ））101は、所定の情報を表示するようになされている。カーソルキー102は、液晶表示部101に表示されたカーソルを移動させるときや、液晶表示部101の表示画面をスクロールさせるときなどに操作される。電源スイッチ104は、ONスイッチ104aおよびOFFスイッチ104bで構成され、それぞれは、装置の電源をオンまたはオフするときに操作される。アルファベットキー105は、アルファベットを入力するときに操作される。決定キー106は、アルファベットキー105を操作することにより入力された文字列を確定したり、また、カーソルキー102を操作

することにより選択した項目を確定したりするときに操作される。ファンクション（機能）キー１０７には、各種の機能が割り当てられており、その機能を実行させるときに操作される。

【００１８】次に、以上のように構成される翻訳装置の動作について説明する。なお、ここでは、翻訳装置は、例えば、英単語を入力すると、その英単語の意味を日本語で記述したもの（英単語を日本語に翻訳したもの）を出力するようになされているものとする。

【００１９】ユーザは、まず最初に、ＯＮスイッチ１０４を操作することにより、装置の電源をオンにする。その後、ユーザは、アルファベットキー１０５を操作することにより、意味を調べようとする英単語を入力する。即ち、例えば、英単語「dog」の意味を調べようとするときは、アルファベットキー１０５のうちの「d」、「o」、「g」の文字に対応するキーを順次操作することで、「dog」を入力する。

【００２０】アルファベットキー１０５を操作することにより入力された文字は、液晶表示部１０１に表示されるようになされており、従って、この場合、液晶表示部１０１には、「dog」が表示される。

【００２１】ユーザは、液晶表示部１０１の表示を見ることで、意味を調べようとする英単語「dog」が正しく入力されたことを確認し、その後、決定キー１０６を操作する。

【００２２】決定キー１０６が操作されると、翻訳装置では、液晶表示部１０１に表示された英単語「dog」の意味（「dog」の翻訳結果）が、その内蔵する英和辞書（電子辞書）（図示せず）から検索され、液晶表示部１０１に表示される。

【００２３】以上のようにして、ユーザにより入力された英単語の意味が表示される。なお、この場合、英単語の意味が、液晶表示部１０１の１画面に表示しきれないときには、カーソルキー１０２を操作して、画面をスクロールさせることにより、その全体を見ることができるようになされている。

【００２４】

【発明が解決しようとする課題】ところで、図１１に示した従来の翻訳装置においては、以下のような課題があった。

【００２５】即ち、日本語から英語、または英語から日本語のうちのいずれの翻訳を行うかを制御部１１に認識させるために、日本人であるユーザＡが発話するとき、およびアメリカ人であるユーザＢが発話するとき、その都度、キースイッチ部１４を操作する必要がある。さらに、そのために、ユーザＡとＢとの間で、翻訳装置を発話をする側に手渡しする必要もあり、このようなことを、発話を行うたびに行うのは面倒であった。

【００２６】そこで、制御部１１に、マイク１Ａに入力された音声の言語を、自動的に識別（自動言語識別）さ

せる方法があるが、言語の識別は困難で、正しい識別結果が得られるとは限らなかった。

【００２７】また、翻訳装置を、ユーザＡおよびＢがそれぞれ携帯していれば、上述の課題を解決することができるが、例えば、ユーザＡが、ユーザＢの国に行く場合において、ユーザＡが、翻訳装置を携帯していくことは考えられるが、ユーザＢが、自国において、翻訳装置を携帯していることは、現実的には考えにくい。

【００２８】そこで、例えば、ユーザＡが、ユーザＢの国に行くときに、翻訳装置を２台携帯していく方法があるが、翻訳装置が携帯型に構成されているといっても、それを２台携帯するのは面倒であり、さらに、この場合、ユーザＡの経済的負担も大になる。

【００２９】一方、図１２に示した従来の翻訳装置においては、以下のような課題があった。

【００３０】即ち、このような翻訳装置を、例えば路上などで使用する場合には、ユーザは、左手または右手のうちのいずれか一方の手で装置を持ち、ＬＣＤ１０１を見ながら、他方の手で装置を操作する必要があった。このため、装置の操作を、ＬＣＤ１０１以外のものを見ながら行うことは、そのものと、ＬＣＤ１０１との間での視線の移動が必要となり、ユーザに煩わしさを感じさせる課題があった。さらに、装置の操作を、筆記用具その他の物を持ちながら行うのは困難であった。

【００３１】そこで、図１１に示した翻訳装置における場合と同様に、装置にマイクおよびスピーカを設け、マイクから意味を調べようとする英単語（日本語に翻訳しようとする英単語）を入力するとともに、その英単語の意味を、スピーカから合成音で出力させるようにする方法がある。

【００３２】しかしながら、装置自体にマイクおよびスピーカを設けた場合には、マイクに混入する雑音を低減するために、ユーザは、装置に設けられたマイクに口を近づけて（あるいは、装置を近づけて）発話を行う必要があり、面倒であった。さらに、環境雑音のレベルが大きい場合には、装置に設けられたスピーカに耳を近づけて（あるいは、装置を近づけて）合成音を聴く必要があり、やはり面倒であった。

【００３３】本発明は、このような状況に鑑みてなされたものであり、翻訳装置の利便性を向上させることができるようにするものである。

【００３４】

【課題を解決するための手段】請求項１に記載の翻訳装置は、音声を入力する入力手段が、複数設けられていることを特徴とする。

【００３５】請求項１０に記載の翻訳装置は、入出力部と本体とが、分離して設けられていることを特徴とする。

【００３６】請求項１に記載の翻訳装置においては、音声を入力する入力手段が、複数設けられている。

【0037】請求項10に記載の翻訳装置においては、入出力部と本体とが、分離して設けられている。

【0038】

【発明の実施の形態】以下に、本発明の実施例を説明するが、その前に、特許請求の範囲に記載の発明の各手段と以下の実施例との対応関係を明らかにするために、各手段の後の括弧内に、対応する実施例（但し、一例）を付加して、本発明の特徴を記述すると、次のようになる。

【0039】即ち、請求項1に記載の翻訳装置は、音声を入力する入力手段（例えば、図1や図4に示すマイク1Aおよび1B、図2に示すマイク24、図3に示すマイク33など）と、入力手段に入力された音声を音声認識する音声認識手段（例えば、図1や図4に示す音声認識部8など）と、音声認識手段の音声認識結果を、入力手段に入力された音声の言語と異なる言語に翻訳する翻訳手段（例えば、図1や図4に示す翻訳部9など）と、翻訳手段の翻訳結果を出力する出力手段（例えば、図1や図4に示すスピーカ2Aおよび2B、図1や図4に示すLCD表示部12、図2に示すスピーカ22、図3に示すスピーカ31など）とを備える翻訳装置であって、入力手段が、複数設けられていることを特徴とする。

【0040】請求項3に記載の翻訳装置は、翻訳手段の翻訳結果に対応する合成音を生成する生成手段（例えば、図1や図4に示す音声合成部10など）をさらに備え、出力手段が、生成手段により生成される合成音を出力することを特徴とする。

【0041】請求項7に記載の翻訳装置は、所定の端末との通信を制御する通信制御手段（例えば、図4に示す電話部43など）をさらに備えることを特徴とする。

【0042】請求項9に記載の翻訳装置は、情報を再生する再生手段（例えば、図4に示すMD（ミニディスク（商標））部45など）をさらに備え、出力手段が、再生手段により再生される情報も出力することを特徴とする。

【0043】請求項10に記載の翻訳装置は、音声を入力する入力手段（例えば、図5乃至図10に示すマイク54など）と、所定の情報に対応する合成音を出力する出力手段（例えば、図5乃至図10に示すスピーカ53など）とを有する入出力部と、入力手段に入力された音声を音声認識する音声認識手段（例えば、図6に示す音声認識回路62など）と、音声認識手段の音声認識結果を、入力手段に入力された音声の言語と異なる言語に翻訳する翻訳手段（例えば、図6に示すCPU65など）と、翻訳手段により出力される翻訳結果に対応する合成音を生成し、出力手段に供給する生成手段（例えば、図6に示す音声合成回路68など）とを有する本体とを備え、入出力部と本体とが、分離して設けられていることを特徴とする。

【0044】なお、勿論この記載は、各手段を上記した

ものに限定することを意味するものではない。

【0045】図1は、本発明を適用した携帯型の翻訳装置の第1実施例の構成を示している。なお、図中、図11における場合と基本的機能が同様の部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。

【0046】即ち、この実施例においては、マイク1Aまたはスピーカ2Aが、スイッチ3の端子3aまたはスイッチ4の端子4aにそれぞれ接続されており、従って、マイク1Aまたはスピーカ2Aは、スイッチ3または4を介して、音声認識部8または音声合成部10とそれぞれ接続されている。また、スイッチ3の端子3bおよびスイッチ4の端子4bは、コネクタ（ヘッドセット用コネクタ）7Aに接続されている。

【0047】さらに、この実施例においては、マイク1Aまたはスピーカ2Aの他のマイクまたはスピーカとして、マイク1Bまたはスピーカ2Bがそれぞれ設けられている。このマイク1Bまたはスピーカ2Bも、マイク1Aまたはスピーカ2Aと同様にして、それぞれ音声認識部8または音声合成部10に接続されている。即ち、マイク1Bまたはスピーカ2Bは、スイッチ5の端子5aまたはスイッチ6の端子6aとそれぞれ接続されており、スイッチ5または6は、音声認識部8または音声合成部10とそれぞれ接続されている。また、スイッチ5の端子5bおよびスイッチ6の端子6bは、コネクタ（ヘッドセット用コネクタ）7Bに接続されている。

【0048】以上のように、音声の入力を行うためのマイクと、合成音の出力を行うためのスピーカとが、複数設けられている。

【0049】コネクタ7Aおよび7Bには、外部のマイクから音声を入力し、また、外部のスピーカに音声を出力するときに、そのマイク、スピーカが接続されるようになされている。なお、コネクタ7Aまたは7Bに、マイクおよびスピーカが外部から接続されたときには、通常は端子3a乃至6aを選択しているスイッチ3乃至6は、端子3b乃至6bを選択するようになされており、これにより、外部のマイクまたはスピーカが、音声認識部8または音声合成部10とそれぞれ電氣的に接続されるようになされている。

【0050】次に、図11における場合と同様に、日本語の話者としてのユーザAと、英語の話者としてのユーザBとが、この翻訳装置を用いて会話を行う場合を例に、その動作について説明する。なお、ここでは、例えば、マイク1Aおよびスピーカ2AをユーザAが、マイク1Bおよびスピーカ2BをユーザBが使用するものとする。

【0051】まず、ユーザAまたはBは、キースイッチ部14を操作し、これにより、マイク1Aおよびスピーカ2Aを、日本語の入出力用として設定するとともに、マイク1Bおよびスピーカ2Bを、英語の入出力用とし

て設定する。このように設定することで、制御部 11 では、以下、マイク 1 A または 1 B には、日本語または英語の音声それぞれ入力されるものとして、また、スピーカ 2 A または 2 B には、日本語または英語の合成音がそれぞれ供給されるように制御が行われる。

【0052】即ち、ユーザ A が日本語による発話を行った場合、その日本語の音声は、マイク 1 A およびスイッチ 3 を介して、音声認識部 8 に供給される。音声認識部 8 は、スイッチ 3 を介して音声を受信すると、その旨を、制御部 11 に出力する。制御部 11 は、音声認識部 8 より、スイッチ 3 を介して音声が入力された旨を受信すると、音声認識部 8 を日本語の音声を音声認識するように、翻訳部 9 を日本語を英語に翻訳するように、音声合成部 10 を英語の合成音を生成するように、それぞれ制御する。

【0053】そして、以下、図 11 における場合と同様の処理が行われ、これにより、LCD 表示部 12 には、マイク 1 A に入力された日本語の音声の音声認識結果、およびその音声認識結果を英語に翻訳した翻訳文が表示され、音声合成部 10 では、その翻訳文に対応する合成音（英語の合成音）が生成される。この翻訳文の合成音は、音声合成部 10 から、スイッチ 6 を介して、スピーカ 2 B に供給されて出力される。これにより、ユーザ B は、スピーカ 2 B から、ユーザ A の発話を英語に翻訳した合成音を聴くことができる。

【0054】一方、ユーザ B が英語による発話を行うと、その英語の音声は、マイク 1 B およびスイッチ 5 を介して、やはり、音声認識部 8 に供給される。音声認識部 8 は、スイッチ 5 を介して音声を受信すると、その旨を、制御部 11 に出力する。制御部 11 は、音声認識部 8 より、スイッチ 5 を介して音声が入力された旨を受信すると、音声認識部 8 を英語の音声を音声認識するように、翻訳部 9 を英語を日本語に翻訳するように、音声合成部 10 を日本語の合成音を生成するように、それぞれ制御する。

【0055】そして、以下、図 11 における場合と同様の処理が行われ、これにより、LCD 表示部 12 には、マイク 1 B に入力された英語の音声の音声認識結果、およびその音声認識結果を日本語に翻訳した翻訳文が表示され、音声合成部 10 では、その翻訳文に対応する合成音（日本語の合成音）が生成される。この翻訳文の合成音は、音声合成部 10 から、スイッチ 4 を介して、スピーカ 2 A に供給されて出力される。これにより、ユーザ A は、スピーカ 2 A から、ユーザ B の発話を日本語に翻訳した合成音を聴くことができる。

【0056】以上のように、マイクが 2 つ設けられているので、最初に、マイク 1 A および 1 B に入力される音声の言語を設定することで、それ以降は、音声が入力されたものかを判別することにより、即ち、ここでは、マイ

ク 1 A または 1 B のうちのいずれかから音声が入力されたかを判別することにより、その音声の言語を、容易に認識することができる。その結果、ユーザ A が発話するときと、ユーザ B が発話するときとで、キースイッチ部 14 を操作する必要がなく、従って、ユーザ A および B に煩わしさを感じさせることを防止することができる。また、そのような操作を防止するために、翻訳装置を 2 台用意する必要もない。

【0057】なお、ユーザ A と B とは、通常は、翻訳装置を挟んで、向き合った状態で会話を行うことが予想されるため、マイク 1 A または 1 B それぞれは、そのような状態で装置が使用されるときに、ユーザ A または B の発話を良好に拾うことができるような指向性を有するものとしてすることができる。また、マイク 1 A および 1 B の指向性は、所定の変更にすることができるようにすることもできる。

【0058】次に、この実施例においては、上述したように、コネクタ 7 A および 7 B が設けられており、このコネクタ 7 A や 7 B には、例えば図 2 に示すヘッドセットや、図 3 に示すイヤークセットなどを接続することができるようになされている。即ち、翻訳装置には、ヘッドセットやイヤークセットなどが外部から接続可能になされている。

【0059】図 2 に示すヘッドセットにおいては、ユーザの頭部に装着可能な装着部 21 にスピーカ 22 が設けられている。なお、スピーカ 22 は、装着部 21 がユーザの頭部に装着されたときに、そのユーザの耳の部分に位置するように取り付けられている。さらに、装着部 22 には、アーム 23 の一端が取り付けられており、その他端には、マイク 24 が設けられている。なお、アーム 23 は、装着部 21 がユーザの頭部に装着されたときに、その先端に設けられたマイク 24 が、ユーザの口の前に位置するように構成されている。

【0060】スピーカ 22 およびマイク 24 には、信号をやりとりするためのコード 25 が接続されており、このコード 25 の、スピーカ 22 およびマイク 24 が接続されていない方の一端には、コネクタ 7 A や 7 B に装着することのできるジャック 26 が設けられている。

【0061】ユーザは、以上のようなヘッドフォンを使用するとき、ジャック 26 をコネクタ 7 A または 7 B に差し込むとともに、装着部 21 を頭部に装着する。

【0062】例えば、ジャック 26 がコネクタ 7 A に差し込まれた場合、スイッチ 3 または 4 は、端子 3 b または 4 b をそれぞれ選択するようになされており、これにより、音声認識部 8 とマイク 24 とが、また、音声合成部 10 とスピーカ 22 とが、それぞれ電氣的に接続される。従って、この場合、ユーザ A が日本語による発話を行うと、その日本語の音声は、マイク 24 およびスイッチ 3 を介して、音声認識部 8 に入力される。また、音声合成部 10 が出力する日本語の合成音は、スイッチ 4 を

介して、スピーカ 22 に供給されて出力される。

【0063】一方、ジャック 26 がコネクタ 7B に差し込まれた場合、スイッチ 5 または 6 は、端子 5b または 6b をそれぞれ選択するようになされており、これにより、やはり、音声認識部 8 とマイク 24 とが、また、音声合成部 10 とスピーカ 22 とが、それぞれ電氣的に接続される。従って、この場合、ユーザ B が英語による発話を行うと、その英語の音声は、マイク 24 およびスイッチ 5 を介して、音声認識部 8 に入力される。また、音声合成部 10 が出力する英語の合成音は、スイッチ 6 を介して、スピーカ 22 に供給されて出力される。

【0064】次に、図 3 に示すイヤークセットにおいては、ユーザの耳に装着可能なスピーカ 31 にアーム 32 が取り付けられている。そして、アーム 32 の、スピーカ 31 が設けられていない用の一端には、マイク 33 が設けられている。なお、アーム 32 は、スピーカ 31 がユーザの耳に装着されたときに、その先端に設けられたマイク 33 が、ユーザの口の前に位置するように構成されている。

【0065】スピーカ 31 およびマイク 33 には、信号をやりとりするためのコード 34 が接続されており、このコード 34 の、スピーカ 31 およびマイク 33 が接続されていない方の一端には、コネクタ 7A や 7B に装着することのできるジャック 35 が設けられている。

【0066】ユーザは、以上のようなイヤークセットを使用するとき、ジャック 35 をコネクタ 7A または 7B に差し込むとともに、スピーカ 31 を耳に装着する。この場合、上述のヘッドセットにおける場合と同様に、マイク 33 に入力された音声は、音声認識部 8 に供給され、また、音声合成部 10 が出力する合成音が、スピーカ 31 に供給される。

【0067】以上のようなヘッドセットやイヤークセットにおいては、マイク 24 や 33 がユーザの口の前に位置するようになされているので、そこには周囲の雑音が混入しにくくなり、従って、音声認識部 8 における音声認識の精度を向上させることが可能となる。さらに、ユーザが、それほど大きな声で発話を行う必要もなくなる。また、ユーザ A または B のうちの少なくともいずれか一方が、ヘッドセットやイヤークセットを使用するようにすることで、一方の音声は、他方が使用するマイクに混入する割合を低減することができる。さらに、スピーカ 22 や 31 は、ユーザの耳の部分に位置するようになされているので、大きな雑音のある環境下であっても、翻訳結果を、明瞭に聴くことが可能となる。

【0068】また、ユーザ A および B の両方が、ヘッドセットやイヤークセットを使用するようにした場合、それぞれの音声を、的確に拾うことが可能となり、その結果、両者の発話が重なっても、音声認識部 8、翻訳部 9、および音声合成部 10 に、その両者の音声に対する処理を、例えば時分割で行わせるようにすることで、そ

のような同時になされた発話に対する対応も可能となる。

【0069】次に、図 4 は、本発明を適用した携帯型の翻訳装置の第 2 実施例の構成を示している。なお、図中、図 1 における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、この翻訳装置は、セレクト 41 および 42、電話部 43、アンテナ (ANT) 44、並びに MD 部 45 が新たに設けられている他は、図 1 の翻訳装置と同様に構成されている。

【0070】セレクト 41 は、制御部 11 により制御され、その入力端子 i1、i2、または i3 に入力される信号のうちのいずれかを選択して、その出力端子 out から出力するようになされている。なお、セレクト 41 の入力端子 i1 乃至 i3 それぞれには、音声合成部 10 が出力する日本語の合成音、MD 部 45 の出力信号、または電話部 43 の出力信号が入力されるようになされており、また、その出力端子 out は、スイッチ 4 と接続されている。

【0071】セレクト 42 は、制御部 11 により制御され、その入力端子 i1 または i2 に入力される信号のうちのいずれかを選択して、その出力端子 out から出力するようになされている。なお、セレクト 42 の入力端子 i1 または i2 それぞれには、マイク 1A が出力する日本語の音声信号（またはコネクタ 7A を介して供給される日本語の音声信号）、または音声合成部 10 が出力する英語の合成音が入力されるようになされており、また、その出力端子 out は、電話部 43 と接続されている。

【0072】電話部 43 は、電話通信を行うための制御を行うようになされている。即ち、電話部 43 は、セレクト 42 の出力端子 out から供給される音声を変調して、アンテナ 44 に出力し、また、アンテナ 44 から供給される信号を復調して音声信号とし、セレクト 41 の入力端子 i3 および音声認識部 8 に出力するようになされている。アンテナ 44 は、電話部 43 からの信号を、電波として放射し、また、図示せぬ電話機（携帯電話機）から基地局を介して送信されてくる電波を受信するようになされている。MD 部 45 は、ミニディスクの再生を行うようになされている。

【0073】次に、その動作について説明する。なお、ここでも、例えば、第 1 実施例における場合と同様に、マイク 1A およびスピーカ 2A を、日本語による発話を行うユーザ A が、マイク 1B およびスピーカ 2B を、英語による発話を行うユーザ B が使用するものとする。

【0074】まず、ユーザ A と B との間で会話を行う場合においては、第 1 実施例における場合と同様に、キースイッチ部 14 を操作する。この場合、制御部 11 は、第 1 実施例における場合と同様の制御を行うとともに、セレクト 41 を、その入力端子 i1 に入力される信号

を、その出力端子outから出力するように制御する。従って、この場合、音声合成部10が出力する日本語の合成音は、セレクト41およびスイッチ4を介して、スピーカ2Aに出力される。

【0075】また、この場合、マイク1Aおよび1Bに入力された音声は、音声認識部8に供給され、さらに、音声合成部10が出力する英語の合成音は、スイッチ6を介してスピーカ2Bに供給される。

【0076】従って、この場合、第1実施例における場合と同様に、ユーザAとBとは会話をを行うことができる。

【0077】次に、ユーザAが、日本語の発話を行う、例えば電話機などの端末を有するユーザ（以下、適宜、電話機日本語ユーザという）と電話通信を行う場合、その旨を指示するように、キースイッチ部14を操作する。この場合、制御部11は、セレクト41を、その入力端子i3に入力される信号を、その出力端子outから出力するように制御するとともに、セレクト42を、その入力端子i1に入力される信号を、その出力端子outから出力するように制御する。

【0078】従って、この場合、マイク1Aに入力されたユーザAの音声は、スイッチ3およびセレクト42を介して、電話部43に供給され、また、電話部43から出力される音声信号は、セレクト41およびスイッチ4を介してスピーカ2Aに供給される。

【0079】さらに、制御部11は、電話部43を電話通信の制御を行うように制御する。これにより、マイク1Aに入力されたユーザAの音声は、スイッチ3およびセレクト42を介して、電話部43に供給され、この電話部43から、電話機日本語ユーザの有する電話機に送信される。一方、電話機日本語ユーザが有する電話機から送信されてきた音声信号は、アンテナ44で受信され、電話機部43に供給される。電話部43は、この音声信号を、セレクト41およびスイッチ4を介して、スピーカ2Aに供給し、これにより、スピーカ2Aからは、電話機日本語ユーザの音声出力される。

【0080】以上のようにして、ユーザAと、電話機日本語ユーザとの間において、電話通信による会話をを行うことができる。

【0081】次に、ユーザAが、英語の発話を行う、例えば電話機などの端末を有するユーザ（以下、適宜、電話機英語ユーザという）と電話通信を行う場合、その旨を指示するように、キースイッチ部14を操作する。この場合、制御部11は、セレクト41を、その入力端子i1に入力される信号を、その出力端子outから出力するように制御するとともに、セレクト42を、その入力端子i2に入力される信号を、その出力端子outから出力するように制御する。さらに、制御部11は、上述の場合と同様に、電話部43を電話通信の制御を行うように制御する。

【0082】この場合、マイク1Aに入力されたユーザAの日本語の音声は、スイッチ3を介して、音声認識部8に供給され、以下、音声認識部8、翻訳部9、および音声合成部10において、第1実施例における場合と同様の処理が行われる。これにより、音声合成部10からは、ユーザAによる日本語の発話を英語に翻訳した翻訳文に対応する合成音出力される。この英語の合成音は、セレクト42において選択されている入力端子i2に供給され、その出力端子outから電話部43に供給される。その結果、ユーザAの発話を英語に翻訳した合成音は、電話機英語ユーザの有する電話機に送信される。

【0083】一方、電話機英語ユーザが有する電話機から送信されてきた音声信号は、アンテナ44で受信され、電話機部43に供給される。電話部43は、この音声信号を、音声認識部8に供給する。この場合、音声認識部8、翻訳部9、および音声合成部10では、第1実施例において、マイク1Bから音声供給された場合と同様の処理が行われる。即ち、この場合、電話機英語ユーザが有する電話機からの音声、つまり英語の音声は、日本語に翻訳され、合成音として、音声合成部10から出力される。この日本語の合成音は、セレクト41において選択されている入力端子i1に供給され、その出力端子outから、スイッチ4を介して、スピーカ2Aに供給される。これにより、スピーカ2Aからは、電話機英語ユーザの音声を日本語に翻訳した合成音出力される。

【0084】以上のようにして、ユーザAと、電話機英語ユーザとの間において、電話通信による会話をを行うことができる。

【0085】次に、ユーザAが、例えばMDに記録された音楽などを聴こうとする場合には、そのMDをMD部45にセットするとともに、そのMDを再生するように、キースイッチ部14を操作する。この場合、制御部11は、セレクト41を、その入力端子i2に入力される信号を、その出力端子outから出力するように制御するとともに、MD部45を、そこにセットされたMDを再生するように制御する。

【0086】これに対応して、MD部45では、MDの再生が行われ、その再生信号が、セレクト41において選択されている入力端子i2に供給され、その出力端子outから出力される。従って、この再生信号は、スイッチ4を介して、スピーカ2Aに供給される。これにより、スピーカ2Aからは、MDから再生された音楽が出力される。

【0087】以上のようにして、ユーザAは、MDから再生された音楽を聴くことができる。

【0088】ところで、この翻訳装置においても、第1実施例における場合と同様に、ヘッドセットやイヤークリップを装着することができるようになされている。上述

のように、MDから再生された音楽を、ヘッドセットやイヤークセットを用いて聴くという翻訳装置の使用態様は、通常の携帯型のラジオ受信機や、カセットテープレコーダなどで、音楽を聴くのと同様の態様であるから、このように、MD部45を設けることによって、翻訳装置においてヘッドセットやイヤークセットを用いることに対する違和感を低減することができる。

【0089】また、第2実施例では、電話部43を設けたことによって、電話による会話が可能となるが、この電話による会話は、ヘッドセットやイヤークセットを用いることにより行い易くなり、さらに、そのような使用態様が定着すれば、やはり、翻訳装置においてヘッドセットやイヤークセットを用いることに対する違和感を低減することができる。

【0090】なお、第2実施例においては、情報を再生する装置として、MD部45を設けるようにしたが、その他、例えばラジオ受信機や、カセットテープレコーダなどを設けるようにすることも可能である。さらに、将来、携帯に便利のように、例えば音楽を記録したICカードなどが実現された場合には、そのようなICカードを再生するための装置を設けるようにすることも可能である。

【0091】また、第1実施例および第2実施例においては、マイクおよびスピーカを2つずつ設けるようにしたが、マイクおよびスピーカは、3以上設けるようにすることも可能である。この場合、3人以上のユーザどうしの間で、異なる言語による会話を行うことが可能となる。

【0092】次に、図5は、本発明を適用した翻訳装置の第3実施例の外観構成を示している。この翻訳装置も、第1実施例および第2実施例における場合と同様に携帯型とされている。

【0093】この翻訳装置は、装置の電源をオン、オフするときに操作される電源スイッチ55が設けられた本体52と、音声を入力するスピーカ53および音声を入力するマイク54とが、独立に、分離して設けられている。即ち、スピーカ53およびマイク54は、本体52自体に設けられているのではなく、本体52とは別に設けられている。なお、図5の実施例では、本体52と、スピーカ53およびマイク54とは、所定の接続線を介して接続されている。

【0094】そして、この翻訳装置においては、マイク54に、例えば日本語の音声を入力すると、それが電気信号としての音声信号に変換され、本体52に供給されるようになされている。本体52では、マイク54に入力された音声は音声認識され、さらに、その音声認識結果が、例えば英語などに翻訳されるようになされている。そして、本体52では、その英語に翻訳された翻訳文に対応する合成音声が生成され、これが、スピーカ53から出力されるようになされている。

【0095】なお、スピーカ53は、ユーザの耳に取り付け可能なように、例えばイヤフォンなどとされている。

【0096】以上のように、スピーカ53は、本体52と分離して設けられ、ユーザの耳に取り付け可能なようになされているので、例えば、環境雑音のレベルが大きい場所において装置を使用する場合であっても、合成音を聞き取ることがことができる。また、マイク54も、本体52と分離して設けられているので、そのマイク54だけを、ユーザの口に近づけるだけで、そこに混入する雑音を低減することができる。

【0097】次に、図6は、図5の翻訳装置の電氣的構成例を示している。マイク54から出力されるアナログの音声信号は、A/D変換器61に供給されるようになされている。A/D変換器61は、マイク54からのアナログの音声信号を、所定のサンプリングクロックにしたがってサンプリングし、さらに量子化することで、デジタルの音声信号に変換するようになされている。A/D変換器61においてデジタル信号とされた音声信号は、音声認識回路62に供給されるようになされている。

【0098】音声認識回路62は、A/D変換器61から音声信号を受信すると、その音声信号の、例えば音程（周波数）や、強弱（振幅）、速度（発話速度）などを音響分析し、その分析結果に基づいて、マイク54に入力された音声は音声認識するようになされている。この音声認識結果は、CPU（Central Processor Unit）65に供給されるようになされている。

【0099】RAM（Random Access Memory）63は、音声認識回路62の動作上必要なデータを記憶するようになされている。ROM（Read Only Memory）64は、音声認識回路62が音声認識の対象とする、例えば単語などを記憶しており、従って、音声認識回路62では、このROM64に記憶されている単語を対象として、マイク54に入力された音声は、音声認識されるようになされている。

【0100】CPU65は、キーインターフェイス72からの信号に対応して、装置の電源を、オン、オフするようになされている。さらに、CPU65は、音声認識回路62および音声合成回路68を制御するようにもなされている。また、CPU65は、音声認識回路62からマイク54に入力された、例えば日本語による音声の音声認識結果を、例えば英語に翻訳するようにもなされている。なお、その結果得られる翻訳文は、音声合成回路68に供給されるようになされている。

【0101】ROM66は、CPU65が行う翻訳に必要なデータが記憶されており、従って、CPU65は、ROM66に記憶されたデータを参照しながら翻訳を行うようになされている。さらに、ROM66には、システムプログラムやアプリケーションプログラムが記憶さ

れており、CPU 65は、このROM 66に記憶されているプログラムを実行することで、上述したような処理を行うようになされている。RAM 67は、CPU 65の動作に必要なデータを一時記憶するようになされている。

【0102】音声合成回路68は、CPU 65から供給される翻訳文に対応した合成音を、ROM 69およびRAM 70にアクセスしながら生成するようになされている。即ち、ROM 69は、例えば、いわゆるCVやVCVなどの単位で音素片データを記憶しており、また、RAM 70は、音声合成回路68の動作に必要なデータを一時記憶するようになされている。そして、音声合成回路68は、CPU 65から供給される翻訳文に対応した合成音（音声データ）を、ROM 69に記憶されている、必要な音素片データを読み出して接続することにより生成するようになされている。

【0103】音声合成回路68で生成された合成音は、D/A変換器71に供給されるようになされている。この合成音は、デジタル信号とされており、D/A変換器71は、このデジタルの合成音をD/A変換することにより、アナログの合成音とするようになされている。このアナログ信号とされた合成音は、スピーカ53に供給され、これにより、スピーカ53からは、マイク54に入力された日本語の音声を英語に翻訳した合成音出力されるようになされている。

【0104】キーインターフェイス72は、電源スイッチ55が操作されると、その操作に対応した信号を、CPU 65に出力するようになされている。

【0105】次に、その動作について説明する。ユーザにより、電源スイッチ55が、装置の電源をオンするように操作されると、キーインターフェイス72から、CPU 65に対して、その操作に対応した信号（以下、適宜、オン信号という）が供給される。CPU 65は、キーインターフェイス72からオン信号を受信すると、装置の電源をオン状態にする。

【0106】その後、マイク54に日本語による音声が入力されると、その音声は、A/D変換器61を介して、音声認識回路62に供給される。音声認識回路62では、マイク54からの音声が入力され、その音声認識結果が、CPU 65に供給される。CPU 65は、音声認識回路62から音声認識結果（日本語の音声認識結果）を受信すると、それを英語に翻訳し、その翻訳文を、音声合成回路68に出力する。音声合成回路68は、CPU 65から翻訳文を受信すると、その翻訳文に対応した合成音を生成し、D/A変換器71を介して、スピーカ53に供給する。これにより、スピーカ53からは、マイク54に入力された日本語の音声を英語に翻訳した合成音出力される。

【0107】次に、以上のような翻訳装置の使用形態（態様）について、図7乃至図10を参照して説明す

る。本体（基体部）52は、図7乃至図10に示すように、ベルト82が取り付け可能になされており、これにより、ユーザは、ベルト82を肩などにかけることにより、本体52を、容易に携帯することができるようになされている。さらに、イヤフォンとして構成されるスピーカ53は、耳に装着することができるようになされている。

【0108】そして、図7の実施例では、マイク54は、例えば、鉛筆やペンなどの筆記具81（筆記具81の一端）に取り付け可能になされている。従って、この場合、ユーザは、例えば、鉛筆やペンなどの筆記具81を持ち、その筆記具81で紙にメモなどを取りながら、翻訳装置を使用することができる。また、翻訳結果は、合成音で、スピーカ53から出力されるので、ユーザは、翻訳結果を見る必要がなく、従って、例えば、紙に書かれた文章などを見ながらであっても、容易に、装置を使用することができる。さらに、この場合、ユーザは、マイク54が取り付けられた筆記具81を口元に移動して発話を行うことにより、マイク54に混入する雑音を低減することができる。

【0109】次に、図8の実施例では、マイク54は、本体52とスピーカ53とを接続する接続線の途中に取り付けられており、これにより、スピーカ53をユーザの耳に装着したときに、マイク54が、ユーザの口元に位置するようになされている。

【0110】図9の実施例では、スピーカ53およびマイク54によりヘッドセットが構成されている。即ち、ユーザの頭部に装着される装着部91には、スピーカ53が設けられている。なお、スピーカ53は、装着部91がユーザの頭部に装着されたときに、そのユーザの耳の部分に位置するように取り付けられている。さらに、装着部91には、アーム92の一端が取り付けられており、その他端には、マイク54が設けられている。なお、アーム92は、装着部91がユーザの頭部に装着されたときに、その先端に設けられたマイク54が、ユーザの口の前に位置するように構成されている。

【0111】図10の実施例では、スピーカ53とマイク54とが、所定の接続線により接続されており、これにより、スピーカ53をユーザの耳に装着したときに、マイク54が、ユーザの口元に位置するようになされている。さらに、この場合、マイク54には、データの送受を無線で行うための通信装置54Aが設けられているとともに、本体52にも、同様の通信装置52Aが設けられており、これにより、本体52と、スピーカ53およびマイク54との間において、信号のやりとりを無線で行うことができるようになされている。

【0112】図8乃至図10の実施例においても、図7の実施例における場合と同様に、スピーカ53が、ユーザの耳に装着されているので、環境雑音のレベルが大きい場所において装置を使用する場合であっても、容易

に、合成音を聞き取ることがことができる。さらに、マイク 5 4 は、スピーカ 5 3 をユーザの耳に取り付けたときに、その口元に位置するようになされているので、マイク 5 4 に混入する雑音を低減することができる。

【0113】なお、第 3 実施例においては、CPU 6 5 に翻訳処理を行われるようにしたが、ROM 6 6 に記憶させるアプリケーションプログラムを変更することによって、CPU 6 5 には、種々の情報処理を行わせるようにすることが可能である。

【0114】さらに、第 3 実施例においては、本体 5 2 と分離したマイクとスピーカとを 1 組だけ設けるようにしたが、このようなマイクとスピーカとは、第 1 実施例および第 2 実施例で説明したように、複数セット設けるようにすることも可能である。

【0115】また、以上の実施例では、日本語と英語との間の翻訳を行うようにしたが、翻訳の対象とする言語は、日本語および英語に限定されるものではない。さらに、翻訳処理は、例えば文単位や、単語単位などの任意の単位で行わせるようにすることが可能である。

【0116】

【発明の効果】請求項 1 に記載の翻訳装置によれば、音声を入力する入力手段が、複数設けられている。従って、複数のユーザは、各入力手段を使用することができるので、装置の利便性を向上させることができる。

【0117】請求項 10 に記載の翻訳装置によれば、入出力部と本体とが、分離して設けられている。従って、入出力部を、ユーザが使用し易い位置に、容易に移動させることが可能となる。

【図面の簡単な説明】

【図 1】本発明を適用した翻訳装置の第 1 実施例の構成

を示すブロック図である。

【図 2】図 1 の翻訳装置に接続可能なヘッドセットを示す図である。

【図 3】図 1 の翻訳装置に接続可能なイヤークセットを示す図である。

【図 4】本発明を適用した翻訳装置の第 2 実施例の構成を示すブロック図である。

【図 5】本発明を適用した翻訳装置の第 3 実施例の外観構成を示す図である。

【図 6】図 5 の翻訳装置の電気的構成例を示すブロック図である。

【図 7】図 5 の翻訳装置の使用状態を示す図である。

【図 8】図 5 の翻訳装置の使用状態を示す図である。

【図 9】図 5 の翻訳装置の使用状態を示す図である。

【図 10】図 5 の翻訳装置の使用状態を示す図である。

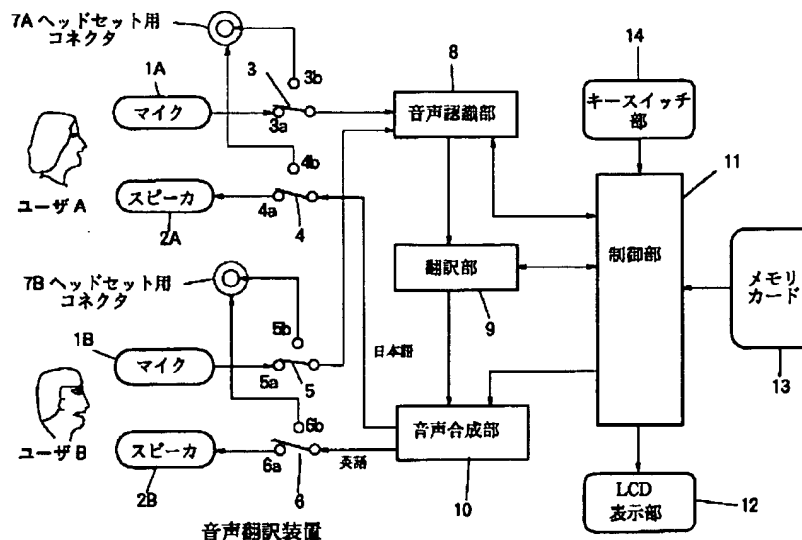
【図 11】従来の翻訳装置の一例の構成を示すブロック図である。

【図 12】従来の翻訳装置の他の一例の構成を示す平面図である。

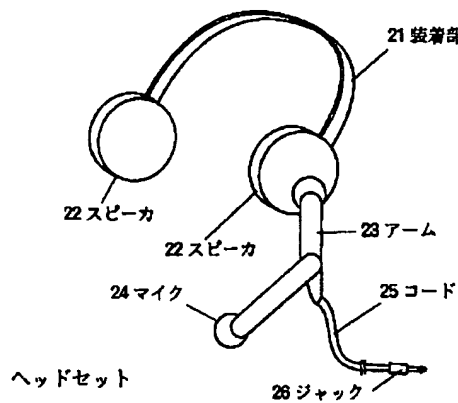
【符号の説明】

1 A, 1 B マイク, 2 A, 2 B スピーカ, 3 乃至 6 スイッチ, 7 A, 7 B コネクタ, 8 音声認識部, 9 翻訳部, 10 音声合成部, 11 制御部, 12 LCD 表示部, 13 メモリカード, 14 キースイッチ部, 22 スピーカ, 24 マイク, 31 スピーカ, 33 マイク, 43 電話部, 44 MD 部, 52 本体, 52 A 通信装置, 53 スピーカ, 54 マイク, 54 A 通信装置, 62 音声認識回路, 65 CPU, 68 音声合成回路, 81 筆記具

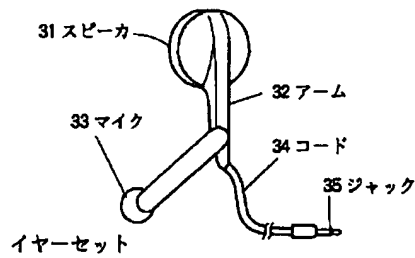
【図 1】



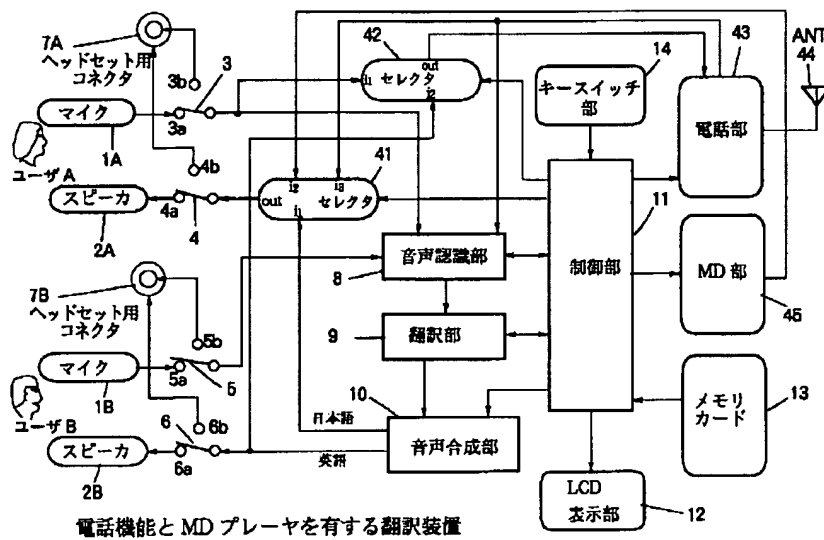
【図 2】



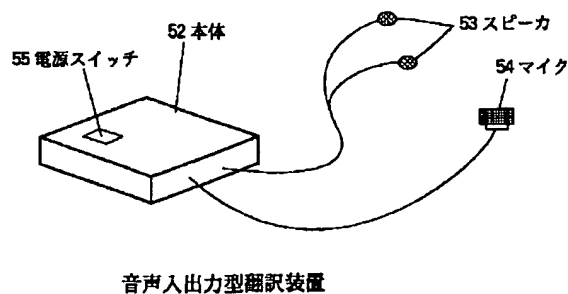
【図 3】



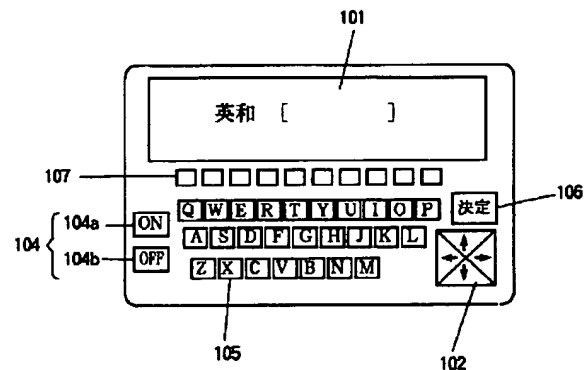
【図 4】



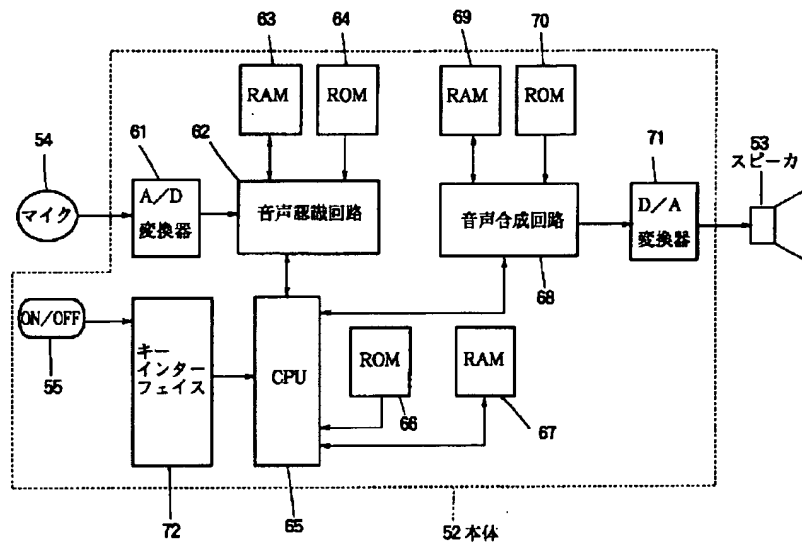
【図 5】



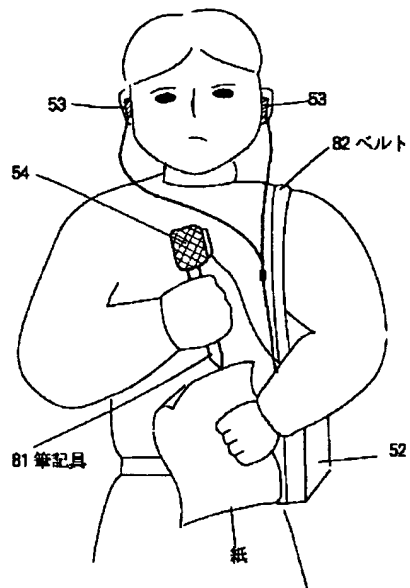
【図 12】



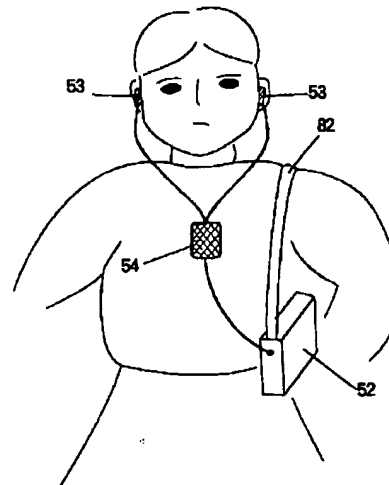
【図 6】



【図 7】

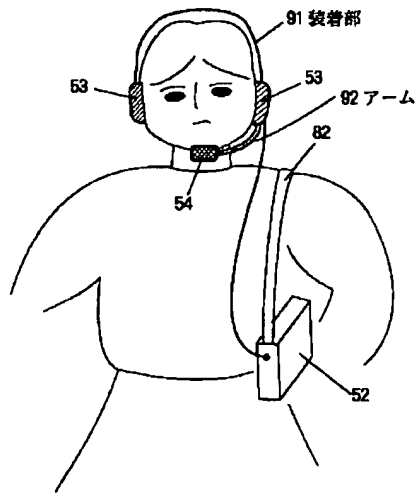


【図 8】

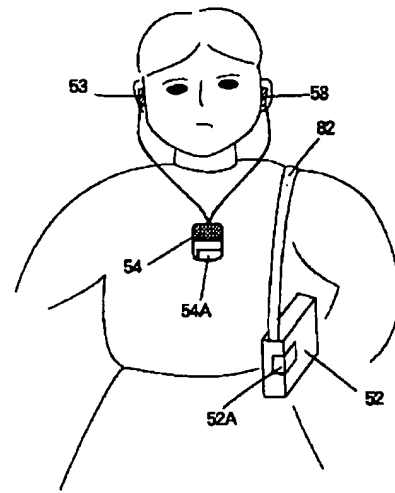


音声入出力型翻訳装置の使用状態

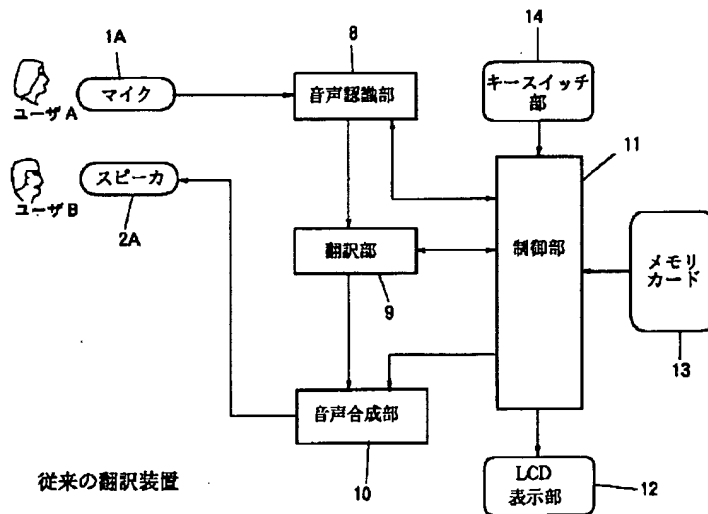
【図 9】



【図 10】



【図 11】



フロントページの続き

(51) Int. Cl. 6
G 1 0 L 3/00

識別記号 5 5 1
庁内整理番号

F I
G 0 6 F 15/20
15/38

技術表示箇所
5 9 2 A
V

11

Japanese Patent Application Kokai Publication Number

Japanese Patent Laid-Open No. H09-292971

Laid-open date 1997/11/11

Application number

Japanese Patent Application No. H08-106746

Filing date 1996/4/26

Applicant Sony

Inventor: Ishii, Kazuo Fujimura, Satoshi [ABSTRACT]

[Problems to be solved] Convenience of a translation device is improved.

[Solution]

When speech production is done in Japanese by user A, the voice goes through microphone 1A, and it is provided in speech recognizer 8. Consider a voice input from microphone 1A to be Japanese, and is done speech recognition in speech recognizer 8, even more particularly, it is English, and is translated with translation part 9. And, as a result, a composite tone corresponding to provided English decodement is generated in voice synthesis part 10, is provided by speaker 2B which user B employs, and is output. On the other hand, the voice goes through microphone 1B, and it is provided in speech recognizer 8 when speech production is done in English by user B. Consider a voice input by microphone 1B to be English, and is done speech recognition in speech recognizer 8, even more particularly, it is Japanese, and is translated with translation part 9. And, as a result, a composite tone corresponding to provided Japanese decodement is generated in voice synthesis part 10, is provided to speaker 2A that user A employs, and is output.

[Detailed description of the invention]

[0001]

[The technical field to which invention belongs]

The present invention relates to translation device (Example: the translation device which is convenient for carrying).

[Problems to be solved by the invention] In conventional translation device shown in FIG. 11, there was the following assignment. [0025] When Japanese user A does speech production to make control section 11 recognize whether our which which is Japanese is translated from English or English from Japanese and when American user B that is to say does speech production, it

was necessary to operate keyswitch part 14 each time. Even more particularly, therefore it is necessary translation device is handed a side doing speech production, and to do it between user A and B, whenever, in such a thing, speech production is done, what was done was troublesome. [0026] Thus there is the approach how identification (automatic speech identification) can put phonic language input into microphone 1A in in control section 11 automatically, but, identification of language was difficult, and right identification result may not have been provided. [0027] In addition, if user A and B carry translation device respectively, the assignment can be solved, but, when, for example, user A goes to a country of user B, it is conceivable that user A carries translation device, but, user B carrying translation device in domestic is hard to be thought about practically. [0028] Thus when, for example, user A goes to a country of user B, there is the approach how two translation device is carried with, but, translation device is configured to a carrying mode, it is troublesome that 2 carry it, and, even more particularly, for this case, it is it to a size charge of economy of user A, though. [0029] On the other hand, in conventional translation device shown in FIG. 12, there was the following assignment. [0030] A user that is to say has device with left hand or a hand of our either of right hand when, for example, such a translation device is employed on streets, while watching LCD 101, it was necessary to operate device with the other hand. Because of this while, in operation of device, watching a thing aside from LCD 101, what is done, move of line of sight at an interval with itself and LCD 101 is necessary, there was the assignment which made a user feel agony. Even more particularly, while, in operation of device, having a thing of writing instrument others, what was done was difficult. [0031] Thus, same as case in translation device shown in FIG. 11, a microphone and loudspeaker are provided to device, alphabetic word (the English word that is going to be translated into Japanese) which is going to examine meaning from a microphone is input, and there is an approach making output meaning of the alphabetic word in composite tone from loudspeaker. [0032] However, it is necessary a user brings a line close to a microphone installed in device (and device is brought close alternatively), and to do speech production to reduce noise to mix with a microphone when a microphone and loudspeaker were provided in device in itself, it was troublesome. Even more particularly, it is necessary ear is brought close to loudspeaker installed in device (and device is brought close alternatively), and to listen to composite tone when an environmental noise level is big, it was also troublesome. [0033] The present

invention was done in view of such a situation, convenience of translation device can be improved. [0034]

[Means to solve the Problems]

As for the translation device as claimed in claim 1, input means to input a voice is characterized by what complement is installed in.

[0035]

It is characterized by what an I/O and a system unit separate a translation device as claimed in claim 10 from it, and is installed in.

[0036]

In a translation device as claimed in claim 1, complement is installed input means to input a voice.

[0037]

An I/O and a system unit separate, and, in a translation device as claimed in claim 10, it is installed.

[0038]

[Embodiment for carrying out the invention]

In the following, example of the present invention is explained.

Example (but, example) coping in a parenthesis after each gateway is added to clarify each gateway and correspondency with the following example of device as claimed in claims, and characteristic of the invention is described first.

[0039]

A translation device as claimed in claim 1 prepares the following.

Input means (microphone 1A / 1B shown in figure 1. FIG. 4, microphone 24 shown in FIG. 2, microphone 33 shown in FIG. 3) to input a voice

A speech recognition gateway (speech recognizer 8 shown in FIG. 1 and FIG. 4) to do speech recognition of a voice input into input means

A translation gateway (compilation department 9 shown in FIG. 1 and FIG. 4) to translate speech recognition result of a speech recognition gateway in the language which is different from phonic language input into input means

Output means (speaker 2A / 2B shown in figure 1.4, LCD display part 12 shown in figure 1.4, speaker 22 shown in FIG. 2, speaker 31 shown in FIG. 3) to output translation result of a translation gateway

It is a translation device comprising the above-mentioned system, and input means is characterized by what complement is installed in.

[0040]

Even more particularly, a translation device as claimed in claim 3 comprises a

generating means (voice synthesis department 10 shown in figure 1.4) to generate a composite tone corresponding to translation result of a translation gateway.

It is characteristic of output means that a composite tone generated by a generating means is output.

[0041]

A translation device as claimed in claim 7 is characterized by that, even more particularly, a communication control gateway (call department 43 shown in FIG. 4) to control communication with a prescribed terminal is comprised.

[0042]

A translation device as claimed in claim 9 comprises a reproducing means (MD (MiniDisk (brand)) part 45 shown in FIG. 4) to reproduce information more.

It is characteristic of output means that information reproduced by a reproducing means is output.

[0043]

A translation device as claimed in claim 10 is characterized by the following.

An I/O having output means (speaker 53 shown in figure 5.10), outputting input means (microphone 54 shown in figure 5.10) to input a voice and a composite tone corresponding to prescribed information.

A translation gateway (CPU 65 shown in FIG. 6) to translate speech recognition result of a speech recognition gateway (speech recognition circuitry 62 shown in FIG. 6) how speech recognition does a voice input into input means and a speech recognition gateway in the language which is different from phonic language input by input means.

A composite tone corresponding to translation result output by a translation gateway is generated, it is a system unit having generating means (electronic speech circuit 68 shown in FIG. 6), to provide in output means.

It is characterized by what the preparation, an I/O and a system unit separate the above-mentioned system from it, and is installed in.

[0044]

In addition, of course it is not a thing meaning that this description limits each gateway to the described above thing.

[0045]

FIG. 1 shows configuration of the first example of a cellular phone which applied the present invention-shaped translation device. It is referred a sign same as case in FIG. 11 about the part that a fundamental facility is similar, and,

therefore, in the following, the description is omitted arbitrarily.

[0046]

In this particular embodiment, microphone 1A / speaker 2A are that is to say connected to switch 3 (terminal 3a) / switch 4 (terminal 4a) respectively.

Thus, microphone 1A / speaker 2A go through switch 3,4, and it is connected to 8. speech recognizer voice synthesis part 10 respectively. Switch 3 (terminal 3b) and switch 4 (terminal 4b) are connected to connector (a connector for head set) 7A.

[0047]

For purposes of this example, as "microphone 1A or a microphone other than speaker 2A or speaker", "microphone 1B / speaker 2B" is installed respectively. "Microphone 1B / speaker 2B" and "microphone 1A / speaker 2A" are connected to speech recognizer 8 or voice synthesis part 10. Microphone 1B or speaker 2B is that is to say connected to terminal 5a of switch 5 or terminal 6a of switch 6 respectively.

Switch 5 or 6 is connected to speech recognizer 8 or voice synthesis department 10 respectively. In addition, terminal 5b of switch 5 and terminal 6b of switch 6 are connected to connector (a connector for head set) 7B.

[0048]

Speaker to output a phonic microphone and composite tone to input, complement mount is installed.

[0049]

It is connected to connector 7A and 7B by outside Mike to input a voice.

In addition, when a voice is output in outside speaker, the microphone speaker is connected to these. When a microphone and speaker were connected to connector 7A / 7B by external, switch 3-6 which, in a typical example, choose terminal 3a — 6a choose terminal 3b — 6b.

An outside microphone or speaker is connected to speech recognizer 8 or voice synthesis department 10 electrically by this respectively.

[0050]

For purposes of example, same as case in FIG. 11, the action is described in the case which user A as a Japanese speaker and user B as an English speaker use this translation device, and talk. Here, user A employs microphone 1A and speaker 2A, user B employs microphone 1B and speaker 2B.

[0051]

User A B operates key switch part 14.

By this, microphone 1A and speaker 2A are set as Japanese input-output

business.

In addition, microphone 1B and speaker 2B are set as English input-output use. As thus described control is done as the thing which Japanese or an English voice is input into in microphone 1A or 1B as follows respectively with control section 11 by setting.

In addition, control is done so that Japanese or an English composite tone is provided in speaker 2A or 2B respectively.

[0052]

When user A did utterance by Japanese, the Japanese voice goes through microphone 1A and switch 3, and it is provided in speech recognizer 8. When speech recognizer 8 receives a voice through switch 3, the effect is output to control section 11. The effect that control section 11 goes through switch 3 from speech recognizer 8, and a voice was input into is received, and this controls as follows afterwards.

"Speech recognition do a voice to be Japanese in speech recognizer 8"

"Translate Japanese into English with compilation department 9"

"Generate a composite tone to be English in voice synthesis department 10"

[0053]

And a process same as case in FIG. 11 is done as follows.

The decodement which translated Japanese phonic speech recognition result input into microphone 1A and the speech recognition result into English is displayed to LCD display part 12 by this.

In voice synthesis department 10, a composite tone (an English composite tone) corresponding to the decodement is generated. A composite tone of this decodement goes through switch 6 from voice synthesis department 10, and it is provided by speaker 2B, and it is output. By this, user B can listen to the composite tone that translated utterance of user A into English from speaker 2B.

[0054]

When user B does utterance by English, the English voice goes through microphone 1B and switch 5, and it is provided in speech recognizer 8. When speech recognizer 8 receives a voice through switch 5, the effect is output to control section 11. When control section 11 goes through switch 5 from speech recognizer 8, and a voice receives the input effect, it controls as follows.

"Speech recognition do a voice to be English in speech recognizer 8"

"Translate English into Japanese with compilation department 9"

"Generate a composite tone to be Japanese in voice synthesis department 10"

[0055]

And a process same as case in FIG. 11 is done as follows.

The decodement which translated English phonic speech recognition result input into microphone 1B and the speech recognition result into Japanese is displayed to LCD display part 12 by this.

In voice synthesis department 10, a composite tone (a Japanese composite tone) corresponding to the decodement is generated. A composite tone of this decodement goes through switch 4 from voice synthesis department 10, and it is provided to speaker 2A, and it is output. By this, user A can listen to the composite tone that translated utterance of user B into Japanese from speaker 2A.

[0056]

Because two microphones are installed, it determines whether a voice goes through our which of switch 3 or 5 by setting phonic language input by microphone 1A and 1B, and was input.

Because a voice determines whether it was input, here, language of the voice can be recognized from microphone 1A or our either of 1B easily. As a result, when user A does utterance, when user B does utterance, it appears, it is not had to operate key switch part 14.

Thus, it can be prevented from making B feel agony.

In addition, it is not had to provide two translation devices to prevent such an operation.

[0057]

With user A and B, that it talks with the status that faced each other is usually expected across a translation device.

When microphone 1A / 1B is such a condition, and a device is employed, the directivity that can pick up speech production of user A B well is comprised. In addition, microphone 1A and directivity of 1B can be changed in prescribed orientation.

[0058]

In this particular embodiment, connector 7A and 7B are comprised as had stated above.

An ear set shown in head set shown in FIG. 2 and FIG. 3 can be connected to this connector 7A and 7B. To a translation device, head set and an ear set are that is to say coupleable from external.

[0059]

In head set shown in FIG. 2, speaker 22 is installed in threaded mounting 21 which can be attached to the head of a user to form a specified shape. When,

speaker 22, threaded mounting 21 was loaded by the head of a user, is installed by aural part of the user to be located. To threaded mounting 22, one end of arm 23 is installed, microphone 24 is installed in the another end. When, arm 23, threaded mounting 21 was loaded by the head of a user, it is configured so that microphone 24 installed in the tip is located before a line of a user.

[0060]

Code 25 to exchange signal is connected to speaker 22 and microphone 24.

Jack 26 which full-fledged, can be attached to connector 7A and 7B to form a specified shape of speaker 22 and this code 25 which microphone 24 is not connected to is installed.

[0061]

When a user employs such a headphone, jack 26 is inserted in connector 7A /7 B, and threaded mounting 21 is attached to head to form a specified shape.

[0062]

When jack 26 was inserted in connector 7A, switch 3/4 choose terminal 3b /4 b respectively.

Speech recognizer 8 and microphone 24 are connected to electricity by this, and, in addition, voice synthesis part 10 and speaker 22 is connected electrically by this. Thus, when in this case user A does utterance by Japanese, a voice of the Japanese goes through microphone 24 and switch 3, and it is input in speech recognizer 8. In addition, the Japanese composite tone that voice synthesis department 10 outputs goes through switch 4, and it is provided to speaker 22, and it is output.

[0063]

When jack 26 was inserted in connector 7B, switch 5/6 choose terminal 5b /6 b.

Speech recognizer 8 and microphone 24 are connected to electricity by this.

In addition, voice synthesis department 10 and speaker 22 is connected electrically by this. When user B does utterance by English, the English voice goes through microphone 24 and switch 5, and it is input in speech recognizer 8. The English composite tone that voice synthesis department 10 outputs goes through switch 6, and it is provided to speaker 22, and it is output.

[0064]

In an ear set shown in FIG. 3, arm 32 is installed to speaker 31 which can be attached to ear of a user to form a specified shape. Microphone 33 is installed in one end of the business that speaker 31 of arm 32 is not installed in. When, arm 32, speaker 31 was loaded by ear of a user, it is configured so that microphone 33 installed in the tip is located before a line of a user.

[0065]

Code 34 to exchange signal is connected to speaker 31 and microphone 33. Jack 35 which full-fledged, can be attached to connector 7A / 7B to form a specified shape of speaker 31 and this code 34 which microphone 33 is not connected to is installed.

[0066]

When the ear set that a user is older than it is employed, jack 35 is inserted in connector 7A / 7 B, and speaker 31 is attached to ear to form a specified shape. For this case, a composite tone and case in the head set and a voice input to microphone 33 similarly are provided in speech recognizer 8, voice synthesis part 10 outputs is provided to speaker 31.

[0067]

In head set such as for example greater or equal and an ear set, it is done so that microphone 24 and 33 are located before a line of a user.

Thus, because ambient noise is hard to become mix there, precision of speech recognition in speech recognizer 8 gets possible to be improved. Even more particularly, a user does not have to do utterance in a so big voice.

Our one of user A / B employs head set and an ear set, and, as thus described, the ratio that one voice mixes with the microphone which the other employs can be reduced. Even if, as for speaker 22.31, there is big noise in what is done to be located in aural part of a user, translation result gets possible to be listened to distinctly.

[0068]

When a thing of user A and B employed head set and an ear set, each voice gets possible to be picked up precisely.

Therefore, even if speech production of both is piled up, speech recognizer 8 and association for the speech production which was done at the same time by making compilation department 9 and voice synthesis part 10 do a process for a voice of the neither in a time sharing are possible.

[0069]

FIG. 4 shows configuration of the second example of a cellular phone which applied the present invention-shaped translation device next. About case in FIG. 1 out of a chart and supporting part, is referred the same sign, in the following, the description is omitted arbitrarily. Selector 41 and 42, telephony part 43, radio antenna (ANT) 44, MD part 45 are that is to say installed in this translation device newly, it is.

Besides it, it is configured same as a translation device of FIG. 1.

[0070]

Selector 41 is controlled by control section 11.

This chooses one of signal input into the input terminal i1, i2, i3, and it is output from the output terminal out. The following is input into input terminal i1 — i3 of selector 41.

These are the Japanese synthesis that voice synthesis department 10 outputs and an output signal of MD part 45 and output signals of telephony part 43.

The output terminal out is connected to switch 4.

[0071]

Selector 42 is controlled by control section 11, and this chooses one of signal input by input terminal i1 / i2, and it is output from the output terminal out. The English composite tone that Japanese audio signal (the Japanese audio signal which connector 7A are gone through, and is provided) which microphone 1A outputs or voice synthesis part 10 outputs is input into input terminal i1 / i2 of selector 42.

The output terminal out is connected to part 43 on the telephone.

[0072]

Part 43 is done on the telephone to do control to do telephone communication. Part 43 modulates a voice provided by output terminal out of selector 42 with it on the telephone, and it is output to radio antenna 44.

In addition, it regains its health, and this does signal provided from radio antenna 44 with audio signal, it is output in input terminal i3 of selector 41 and speech recognizer 8. Radio antenna 44 emits signal from part 43 as radio wave on the telephone, the radio wave that a base station is gone through, and have been transmitted by telephone (portable telephone) which is not illustrated is received. MD part 45 is done to play a MiniDisk.

[0073]

The action is explained next. By way of example only, same as case in the first example, it is assumed that user A doing speech production by Japanese employs microphone 1A and speaker 2A, and user B doing speech production by English employs microphone 1B and speaker 2B.

[0074]

When it talks between user A and B, same as case in the first example, key switch part 14 is operated. For this case, control section 11 does control same as case in the first example, and this controls selector 41 to output signal input by input terminal i1 from output terminal out.

Thus, the Japanese composite tone that in this case voice synthesis department

10 outputs goes through selector 41 and switch 4, and it is output to speaker 2A.
[0075]

For this case, a voice input into microphone 1A and 1B is provided in speech recognizer 8.

The English composite tone that voice synthesis department 10 outputs goes through switch 6, and it is provided by speaker 2B.

[0076]

Thus, in this case it can talk with user A and B same as case in the first example.

[0077]

When user A does Japanese utterance (user (telephone Japanese user) having a terminal such as telephone and telephone communication are done), key switch part 14 is operated to order the effect. For this case, control section 11 controls to output the signal that it is input selector 41 into the input terminal i3 from the output terminal out.

In addition, it controls to output signal that is this controls selector 42 as follows, it is input by input terminal i1 from the output terminal out.

[0078]

Thus, a voice of user A which in this case was input into microphone 1A goes through switch 3 and selector 42, and it is provided to telephony part 43.

Audio signal output from part 43 on the telephone goes through selector 41 and switch 4, and it is provided to speaker 2A.

[0079]

Even more particularly, control section 11 controls part 43 on the telephone to control telephone communication. A voice of user A input into microphone 1A by this goes through switch 3 and selector 42, and it is provided to telephony part 43.

Therefore, this is transmitted to the telephone which a telephone Japanese user comprises from part 43 on the telephone. The audio signal which has been transmitted by the telephone which a telephone Japanese user comprises is received with radio antenna 44, is provided to telephone part 43. Selector 41 and switch 4 is gone through, and part 43 provides this audio signal to speaker 2A on the telephone, a voice of a telephone Japanese user is output from speaker 2A by this.

[0080]

It is done as above, and, in user A and an interval with a telephone Japanese user, dialog by telephone communication can be done.

[0081]

When user A does English utterance (telephone communication is done with user "telephone English user" having a terminal such as telephone), key switch part 14 is operated to order the effect. For this case, control section 11 controls to output the signal that it is input selector 41 into the input terminal i1 from the output terminal out.

In addition, this controls selector 42 as follows.

Signal input into input terminal i2 is output from the output terminal out. Control section 11 controls part 43 same as the case on the telephone to control telephone communication.

[0082]

For this case, a Japanese voice of user A input into microphone 1A goes through switch 3, and it is provided in speech recognizer 8.

In speech recognizer 8, compilation department 9, voice synthesis department 10, a process same as case in the first example is done. A composite tone corresponding to the decodement which translated Japanese utterance by user A into English is output from voice synthesis department 10 by this. This English composite tone is provided by input terminal i2 chosen in selector 42, is provided to telephony part 43 by the output terminal out. As a result, the composite tone that translated utterance of user A into English is transmitted to the telephone which telephone English user comprises.

[0083]

On the other hand, the audio signal which has been transmitted by the telephone which telephone English user comprises is received with radio antenna 44, is provided to telephone part 43. Part 43 provides this audio signal in speech recognizer 8 on the telephone. For this case, a process same as the case that a voice was provided in the first example in speech recognizer 8, compilation department 9 and voice synthesis part 10 by microphone 1B is done. For this case, it is Japanese, and a voice (an English voice) from the telephone which telephone English user comprises is translated, as a composite tone, is output by voice synthesis part 10. This Japanese composite tone is provided by input terminal i1 chosen in selector 41, switch 4 is gone through, and is provided to speaker 2A by the output terminal out. The composite tone that translated a voice of telephone English user into Japanese is output from speaker 2A by this.

[0084]

It is done as above, and, in user A and an interval with telephone English user,

dialog by telephone communication can be done.

[0085]

Key switch part 14 is operated the MD is set at MD part 45 when, for example, user A is going to listen to recorded music to an MD, and to reproduce the MD. For this case, control section 11 controls to output the signal that it is input selector 41 into the input terminal i2 from the output terminal out, and MD part 45 is controlled to reproduce an MD set there.

[0086]

In correspondence with this, reconstruction of an MD is done with MD part 45, the regenerative signal is provided by input terminal i2 chosen in selector 41, is output by the output terminal out. Thus, this regenerative signal is provided to speaker 2A through switch 4. By this, reproduced music is output from speaker 2A by an MD.

[0087]

It is done as above, and user A can listen to reproduced music from an MD.

[with 0088] it is about time it, and, in this translation device, it is done same as case in the first example to be able to attach head set and an ear set to form a specified shape.

Like statement above, use aspect of a translation device head set and an ear set are used, and to listen to reproduced music from an MD is aspect same as what listen to music in these (normal cellular phone-shaped radio broadcast receiver, cassette recorder).

As thus described, by means of providing MD part 45, sense of incongruity for using head set and an ear set in a translation device can be reduced.

[0089]

In addition, in the second example, part 43 is provided on the telephone, and dialog with telephony is enabled by callosity, but, dialog with this telephony is easy to become do by using head set and an ear set.

Even more particularly, if such a use aspect is firmly established, after all sense of incongruity for using head set and an ear set in a translation device can be reduced.

[0090]

In addition, in the second example, MD region 45 was provided as the device which reproduced information.

By way of example only, radio broadcast receiver or cassette recorder can be made. Even more particularly, a device to reproduce such an IC card can be

provided when the IC cards which recorded music were realized in the future to be convenient for carrying.

[0091]

In addition, in the first example and the second example, a microphone and two speaker were provided.

A microphone and speaker can be made more than 3. For this case, at an interval between users more than three people, dialog by different language gets possible to be done.

[0092]

FIG. 5 shows appearance configuration of tertiary example of the translation device which applied the present invention next. This the first translation device is done with a carrying model same as example and case in the second example.

[0093]

This translation device turns on an electric power supply of a device, when it turns off, microphone 54 inputting speaker 53 outputting system unit 52 which operated power switch 55 was installed in and a voice and a voice separates independency from it, and it is installed. Speaker 53 and microphone 54 is not that is to say installed in system unit 52 in itself.

It is installed separately from system unit 52. In addition, prescribed interconnect line is gone through, and, in example of FIG. 5, it is connected to system unit 52 and speaker 53 and microphone 54.

[0094]

And it is converted by audio signal as electrical signal when, for example, it inputs in a Japanese voice to microphone 54 in this translation device, it is done so that is provided to system unit 52. Speech recognition leave, and, even more particularly, a voice input into microphone 54 with system unit 52 does it so that, for example, the speech recognition result is translated by English. And, with system unit 52, a composite tone corresponding to a translated sentence translated into the English is generated, it is done so that this is output from speaker 53.

[0095]

In addition, for example, speaker 53 is considered to be earphones in ear of a user to be mountable.

[0096]

Speaker 53 divides it with system unit 52 as above, and it is installed.

It is done in ear of a user to be mountable.

Even if an environmental noise level employs a device in big location, it is possible for a thing to hear a composite tone. In addition, because microphone 54 separates with system unit 52, too, and it is installed, only the microphone 54 is brought close to a line of a user, noise to mix with there can be reduced.

[0097]

FIG. 6 shows electric configuration example of a translation device of FIG. 5 next. Analog audio signal output from microphone 54 is done so that it is provided to A/D converter 61. A/D converter 61 wants to do analog audio signal from microphone 54 to a prescribed sampling clock, and it is sampled, it is done to convert it for digital audio signal in what, even more particularly, is quantized. Audio signal considered to be a digital signal in A/D converter 61 is done so that it is provided to speech recognition circuitry 62.

[0098]

When speech recognition circuitry 62 receives audio signal from A/D converter 61, an acoustic analysis does the audio signal (musical interval (frequency) and strength and weakness (swing), velocity (speech production velocity)).

Based on the analyses result, a voice input into microphone 54 is done to do speech recognition. This speech recognition result is done so that it is provided to CPU (Central Processor Unit) 65.

[0099]

RAM (Random Access Memory) 63 is done to store necessary data in action of speech recognition circuitry 62. ROM (Read Only Memory) 64, speech recognition circuitry 62, of speech recognition, it is intended for.

By way of example only, words are memorized.

Thus, a voice input to microphone 54 as object does a word stored to this ROM 64 with speech recognition circuitry 62 so that it is done speech recognition.

[0100]

CPU 65 supports signal from key interface 72, and it is done so that on turns off an electric power supply of a device. Even more particularly, CPU 65 is done to control speech recognition circuitry 62 and electronic speech circuit 68. In addition, CPU 65 was input into microphone 54 from speech recognition circuitry 62, it is done, for example, for example, it is English, and to translate phonic speech recognition result by Japanese. In addition, as a result, provided decodement is done so that it is provided to electronic speech circuit 68.

[0101]

As for ROM 66, data to be necessary for the compilation that CPU 65 does is

stored.

Thus, CPU 65 is done to translate it when taken in conjunction with data memorized to ROM 66. Even more particularly, to ROM 66, a system program and application program are memorized.

It is done to work to have stated CPU 65 above by executing a program memorized to this ROM 66. RAM 67 is done to store necessary data in action of CPU 65 at one time.

[0102]

While accessing RAM 70 that ROM 69 invites a composite tone corresponding to the decodement that electronic speech circuit 68 is provided from CPU 65, it is done to generate it. ROM 69 that is to say stores phoneme piece data in a unit such as CV and VCV.

In addition, RAM 70 is done to store necessary data in action of electronic speech circuit 68 at one time. Electronic speech circuit 68 generates a composite tone (audio data) corresponding to decodement provided from CPU 65 to ROM 69 by the stored phoneme piece data that it is necessary is read, and connecting it.

[0103]

A generated composite tone is done with electronic speech circuit 68 so that it is provided to D/A converter 71. This composite tone is considered to be a digital signal, D/A converter 71 is done to do this digital composite tone with an analog composite tone by D/Aing convert. A composite tone considered to be this analog signal is provided to speaker 53.

A Japanese voice input into microphone 54 from speaker 53 by this is done so that a translated composite tone is output by English.

[0104]

When power switch 55 is operated, key interface 72 is done to output signal corresponding to the operation to CPU 65.

[0105]

The action is explained next. Signal (an ON signal is said) corresponding to the operation is provided for CPU 65 from key interface 72 by a user when operated so that power switch 55 turns on an electric power supply of a device. When CPU 65 receives an ON signal from key interface 72, an electric power supply of a device is done in an ON state.

[0106]

When a voice by Japanese is input into microphone 54, the voice goes through A/D converter 61, and it is provided to speech recognition circuitry 62

afterwards. With speech recognition circuitry 62, a voice from microphone 54 leave speech recognition, and the speech recognition result is provided to CPU 65. When CPU 65 receives speech recognition result (Japanese speech recognition result) from speech recognition circuitry 62, it is English, and it is translated, the decodement is output to electronic speech circuit 68. When electronic speech circuit 68 receives decodement from CPU 65, a composite tone corresponding to the decodement is generated, D/A converter 71 is gone through, and it is provided to speaker 53. The composite tone that translated a Japanese voice input into microphone 54 into English is output from speaker 53 by this.

[0107]

Use configuration of a translation device such as for example greater or equal is described referring to figure 7-10 next. Belt 82 is mountable, and system unit (a base member) 52 is done as shown in figure 7-10.

By this, a user does it to be able to carry system unit 52 easily by hanging belt 82 to shoulders. Even more particularly, speaker 53 configured as an earphone is done to be able to attach it to ear to form a specified shape.

[0108]

And, for example, it is mountable, and, in example of FIG. 7, microphone 54 is done to writing implements 81 (one end of writing implements 81) such as pencil and pen. Thus, in this case, for example, a user has writing implements 81 such as pencil and pen, while taking helpful information as paper with the writing implements 81, a translation device can be employed. In addition, because translation result is a composite tone, and is output from speaker 53, a user does not have to watch translation result.

Thus, while watching sentence written to paper, even if you are met, a device can be employed easily. Even more particularly, in this case a user can reduce noise to mix writing implements 81 that microphone 54 was installed with microphone 54 by mouth is moved to, and doing utterance.

[0109]

In example of FIG. 8, microphone 54 is installed in the middle of the interconnect line which connects speaker 53 to system unit 52 next.

When, by this, speaker 53 was attached to ear of a user to form a specified shape, it is done so that microphone 54 is located in mouth of a user.

[0110]

In example of FIG. 9, head set is configured with speaker 53 and microphone 54. Speaker 53 is that is to say installed in threaded mounting 91 loaded by the

head of a user. In addition, when, speaker 53, threaded mounting 91 was loaded by the head of a user, is installed by aural part of the user to be located. Even more particularly, to threaded mounting 91, one end of arm 92 is installed, microphone 54 is installed in the another end. In addition, when, arm 92, threaded mounting 91 was loaded by the head of a user, it is configured so that microphone 54 installed in the tip is located before a line of a user.

[0111]

In example of Fig.10, speaker 53 and microphone 54 are connected by prescribed interconnect line.

When, by this, speaker 53 was attached to ear of a user to form a specified shape, it is done so that microphone 54 is located in mouth of a user. Even more particularly, in this case communications device 54A to do transmission and reception of data by radio are installed in microphone 54.

Communications device 52A are installed in system unit 52.

It is done to be able to do the exchanges of signal by this by radio in system unit 52 and an interval with speaker 53 and microphone 54.

[0112]

In example of figure 8-10, speaker 53 is attached to ear of a user to form a specified shape same as case in example of FIG. 7.

Therefore, even if an environmental noise level employs a device in big location, it is possible for a thing to hear a composite tone easily. Even more particularly, when microphone 54 installed speaker 53 in ear of a user, it is done to be located in the mouth.

Therefore, the noise that microphone 54 is contaminated with can be reduced.

[0113]

In tertiary example, it was translated to CPU 65.

By means of changing application program making ROM 66 store it, it can make CPU 65 do various kinds of data processing.

[0114]

Even more particularly, in tertiary example, only lot provided system unit 52 and separate microphone and speaker.

A complement set can be provided as having described it in the first example and the second example with such a microphone and speaker.

[0115]

In addition, in the above-mentioned example, Japanese and compilation between things of English were done.

The language which translation intends for is not a thing limited to Japanese

and English. Even more particularly, it can make, for example, do translation in an optional unit such as a statement unit and a word unit.

[0116]

[Effect of the Invention]

According to a translation device as claimed in claim 1, complement is installed input means to input a voice. Thus, an a plurality of user can improve convenience of a device in what can employ each input means.

[0117]

An I/O and a system unit separate it from it, and, according to a translation device as claimed in claim 10, it is installed. Thus, the position which a user is easy to employ gets possible to move an I/O easily.

Brief Description of the Drawings

[FIG. 1] It is a block diagram to show configuration of the first example of the translation device which applied the present invention in.

[FIG. 2] It is a chart to show the head set which is accessible to a translation device of FIG. 1 in.

[FIG. 3] It is a chart to show the ear set which is accessible to a translation device of FIG. 1 in.

[FIG. 4] It is a block diagram to show configuration of the second example of the translation device which applied the present invention in.

[FIG. 5] It is a chart to show appearance configuration of tertiary example of the translation device which applied the present invention in.

[FIG. 6] It is a block diagram to show electric configuration example in of a translation device of FIG. 5.

[FIG. 7] It is a chart showing busy condition of a translation device of FIG. 5.

[FIG. 8] It is a chart showing busy condition of a translation device of FIG. 5.

[FIG. 9] It is a chart showing busy condition of a translation device of FIG. 5.

[Fig.10]

It is a chart showing busy condition of a translation device of FIG. 5.

[FIG. 11]

It is a block diagram to show configuration of an example of a conventional translation device in.

[FIG. 12]

It is a plan view to show configuration of an example other than a conventional translation device in.

[Denotation of Reference Numerals]

(1A, 1B) a microphone,

(2A, 2B) speaker,
(3-6) it is switched to,
(7A, 7B) a connector,
(8) a speech recognizer,
(9) compilation department,
(10) a voice synthesis department,
(11) control section,
(12) LCD display part,
(13) a memory card,
(14) key switch part,
22 speaker,
24 microphones,
31 speaker
33 microphones
43 calls department,
44 MDs part,
52 system units,
52A communications device,
53 speaker,
54 microphones,
54A communications device,
62 speech recognition circuitries,
65 CPU,
68 electronic speech circuit,
81 writing implements